

Digitale Tontechnik im Studio

Grundfragen eines Standards für digitale Tonsignale

Digital-Schallplatten, digitale Tonaufzeichnung und Übertragung gehören zweifellos zur Entwicklung der Audiotechnik in den 80er Jahren. Doch bevor die Endprodukte vielleicht als „Super-HiFi“ auf den Markt kommen, sind wichtige und nicht ganz einfache Normungsfragen zu klären. Sie waren das erste Thema einer Diskussions-Sitzung der Nachrichtentechnischen Gesellschaft NTG im Berliner Heinrich-Hertz-Institut, während der auch noch die Analog/Digital- und Digital/Analog-Wandlung sowie Aufzeichnung und Bearbeitung behandelt wurden. Hier das Wichtigste aus dem Beitrag von Björn Blüthgen, Poylgram Hannover, über Grundfragen zur Festlegung eines Standards.

Digitale Tontechnik im Studio betrifft drei Hauptanwendungsbereiche:

- den Einsatz zur operationellen Unterstützung des Studioperpersonals durch elektronische Abmischcomputer oder Routingfelder,
- die Anwendung in Peripheriegeräten zur Nachhallerzeugung und für später erwartete rein digitale Mischpulte und
- die digitale Tonaufzeichnung. Die Erkenntnis der Notwendigkeit von weltweit kompatibelem Programmträgeraustausch machte die digitale Tonaufzeichnung mit den daran beteiligten Parametern, wie Abtastfrequenz und Fehler-Korrekturverfahren, zum Hauptthema.

Viele Beteiligte gingen dabei von der Idealvorstellung einer Standardisierung nach vorheriger Absprache anhand teilweise theoretischer Voraussetzungen aus, wie zum Beispiel die Diskussion um die Abtastfrequenz deutlich zeigte.

Doch ein Vorschlag zur Veröffentlichung als Standardempfehlung über das nationale Normungsgremium dem IEC¹⁾ vorgelegt, muß sich jedoch auf eine realisierte Geräteausführung beziehen. Von mehr als einem Land eingereichte unterschiedliche Vorschläge zum gleichen Anwendungsfall können dann durch die IEC harmonisiert werden. Die AES²⁾ kann dabei als Design Committee oder Plattform für Gespräche zwischen Hersteller und Anwender fungieren.

Eine Untersuchung von existierenden Quellencodier-Charakteristiken zeigt nach einem Auszug aus dem Sekretariatsbericht der IEC Working Group 15, Digital Audio Recording, daß eine gewisse Praxis besteht, und daß zwei schon in Gebrauch befindliche sowie zwei in der Entwicklungsphase stehende Recorder

1. eine Abtastfrequenz von 50 kHz
 2. für die Quantisierung - 16 bit linear
 3. den „Zweier Komplementär Binär Code“ und
 4. keine Quellencodier-Entzerrung anwenden oder vorsehen.
- Damit sind 50 kHz aber noch nicht als Standard empfohlen!

Diskussionen mit renommierten Recorder-Herstellern und Studios mit täglicher Vielspurrecorder-Erfahrung ergibt etwa folgendes Bild

Ampex hat sein früher veröffentlichtes Konzept zur Datenformatierung verworfen und will Ende 1980 eine verbesserte Version in Stereo-Ausführung auf den Markt bringen. Man sieht dort die Rundfunkstationen als einen ergiebigeren Markt (gegenüber nur wenigen teuren Multitrackrecordern für einzelne professionelle Studios).

MCI baut in Kooperation für EMI, England, eine größere Auftragsserie, wenn die Tests anhand von 2 ausgelieferten Prototypen bei EMI, London, abgeschlossen sind. Auch dies sind Stereo-Recorder. EMI beabsichtigt damit, die ursprüngliche Qualität von analogen Multitrack-Abmischungen auf Digital-Stereo in allen Produktionsstufen weltweit zu erhalten.

Aus der Codierungspraxis: 50 kHz und linear mit 16 bit codiert.

3M mit je einem 32- und 4-Spur-digital-Recorder in bereits 5 Studios (die Zahl wird schnell zunehmen) bei täglichem Betrieb, will seinen „Standard“ durch Fakten durchsetzen. Der früher mit Bildschirmterminal arbeitende Prototyp für elektronisches Editing wurde wegen nicht praxistgerechter Arbeitsweise verworfen und dafür ein vorerst einfacher Editor ohne weiches „Crossfade“³⁾ bei Signal-Einblendungen vorgestellt. Random Punch-Ins⁴⁾ an tieffrequenten Signalen sind teilweise leicht hörbar (40 %), jedoch können unhörbare Schnitte durchgeführt werden. Ein im Crossfade weiter verbesserter Editor⁵⁾ wird für Mai 1980 erwartet.

Ein Digital-Audio-Stereo-Recorder, der für sich beansprucht, neben dem elektronischen Schnitt auch mechanische Schnitte und Aufnahmen auf Bändern mit Klebestellen zuzulassen, wurde von Mitsubishi vorgestellt. Hier wird trickreich die zeitliche Koinzidenz von Dropouts auf mehreren Datenspurten sowie auf der Kontrollspur für weiche, interpolativ abgeleitete Signalüberblendungen ausgenutzt. Die Stärke des Fehler-Korrekturverfahrens liegt in der Anzahl von 4 Datenspurten pro Audiosignal, wobei die Sample-Segmente zusätzlich zeitlich verschachtelt (interleaved) sind.

³⁾ Crossfade = (Schnittstellen-)Überblendung.

⁴⁾ Random Punch-Ins = Programm-Einblendungen an beliebiger Stelle.

⁵⁾ Editor = Schnittstellen-Prozessor.

¹⁾ IEC = International Electrotechnical Commission.

²⁾ AES = Audio Engineering Society.

Man wird kaum eine derartige Technik auch für Multitrack-Ausführungen erwarten können, da die ökonomische Fertigung von zum Beispiel 96-Spur-Kopfträgern für 24-Spur-Recorder erhebliche Schwierigkeiten bereiten wird. Auch die Problemlösung mittels Dünnschicht-Technik ist fraglich, wenn von glaubwürdiger Seite 6/10 Ampere pro Impuls, also etwa 58 A für 24 Audio-Kanäle, als erforderlich bezeichnet werden (Störstrahlung!).

Sony überraschte mit einem tragbaren 4-Kanal-Recorder für 1/4-Zoll-Band-Material, dessen mechanisches Konzept auch für eine 24-Spur-1-Zoll-Ausführung gelten soll. Während der 1/4-Zoll-Typ mit 2 Datenspuren pro Audiosignal eine Bandgeschwindigkeit von nur 38 cm/sec benötigt, werden die Multitrack-Recorder je Audiosignal nur eine Datenspur erhalten, dabei jedoch 76 cm/sec Band bei einer nominalen Abtastfrequenz von 50,4 kHz verbrauchen.

Die Lösung hängt von den Bandpässen im Codec ab.

Gespräche mit Anwendern über Erfahrungen an den verschiedenen Recordern in der Praxis (3-M. Sony und Soundstream) bestätigen die an die Digitaltechnik gestellten Erwartungen hinsichtlich Klangverbesserung, wobei aber vereinzelt geäußerte subjektiv unterschiedliche Bevorzugung der verschiedenen Codecs⁶⁾ eine nähere Untersuchung der Ursachen empfiehlt.

Interessant war auch über die Qualitätsverbesserung bei 16 Bit Linear-Quantisierung zu hören, wenn der zu übertragende Abtastwert mittels Integration aus einem Vielfachen der Abtastfrequenz auf die 18 Bit entsprechende Qualität gebracht wird.

Für die Harmonisierung der Grundparameter für eine internationale Bandaustauschbarkeit soll ein von Anwendern den Herstellern vorgelegter Fragenkatalog als Arbeitsgrundlage dienen. Der Fragenkomplex für die Austauschbarkeit der Bänder betrifft:

- a) die freie Austauschbarkeit bespielter Bänder
- b) den Bandverbrauch
- c) den Programm-Schnitt, die klangtechnische Bearbeitung sowie Inserttechnik
- d) den Recorderabgleich und erforderliche Servicearbeiten
- e) die Schnittstellenbedingungen für den Anschluß peripherer Geräte
- f) das Speichern von Tonprogramm-Kontrollsignalen.

Dreimal Digitalaudio für Rundfunk-Übertragungsketten, kompatibel mit dem Fernsehton und als Tonkonserve für den Endverbraucher.

Neben der Frage der Bandaustauschbarkeit sind weitere von grundsätzlicher Bedeutung und für eine Harmonisierung unter wirtschaftlichen Gesichtspunkten offen. Diese beziehen sich auf Hörbarkeitsgrenzen, den Stand der Technik, die Gerätekosten sowie Umwelteinflüsse, wie maximal/minimal zulässige Umgebungstemperaturen, Lüftergeräusche und elektrische Stör(ein-)strahlungen.

Die Wahl der Abtastfrequenz als ein Hauptdiskussionspunkt (ein Sony-Prototyp verarbeitet 7 verschiedene Frequenzen) sollte somit auch logischen Gesichtspunkten entsprechen und nicht nach Energie verschwendenden, einzeln aus dem Zusammenhang gezogenen, nicht stichhaltigen Werbegewohnheiten erfolgen. In der digitalen Ebene eines Programmkanals ist zum Beispiel nicht mit der Zunahme von Gruppenlaufzeitverzerrungen in der Produktionskette zu rechnen. Wie erneute detaillierte Untersuchun-

gen von JVC (siehe AES-New York 79) wiederum bestätigten, sind 15 kHz Hörbandbreite für abgemischtes Programm-Material völlig ausreichend.

Für welche Anwendungszwecke sollte eine höhere Abtastfrequenz dennoch möglich sein?

- 1.) Beim Hintereinanderschalten von mehreren Codecs, aber wieviel Codecs sollen mindestens ohne signifikante Qualitäts-einbußen hintereinander schaltbar sein?
- 2.) Für „Over-Dubbing“⁷⁾ wie für die mehrstimmige Aufnahme eines Sängers mittels reduzierter Wiedergabegeschwindigkeit und anschließender „Synchron“-Aufnahme mit verringerter Abtastfrequenz, werden bis zu -10 % der Nominalgeschwindigkeit ohne hörbare Aliasing-Produkte⁸⁾ gefordert.

Die Antworten hierzu sind abhängig von der praxisgerechten Lösung der im Codec notwendigen Bandpassprobleme.

Das Eingangs-Tiefpassfilter muß spektrale Tonfrequenzanteile und zum Beispiel Übersprechkomponenten von peripheren Geräten mit abweichender Abtastrate oberhalb der Hörbandbreite ausreichend unterdrücken ebenso wie auch Ausgangs-Tiefpassfilter das Entstehen von Aliasing-Produkten vermeiden sollen. Eine Apertur-Korrekturentzerrung ist ebenfalls zur Kompensation der den Abtast- und Halteschaltungen typischen Bandpasseigenschaft erforderlich. Weiterhin können für einige Anwendungsfälle Pre-/Deemphasis-Filter vorteilhaft sein.

Zu den Grundfragen aus dem Bereich der Filterprobleme gehört die nach einem Untersuchungsbericht über die Gründe der Forderung nach 20 kHz ohne relevante hörbare Eigenschaften. Auch ist zu fragen, welche technischen, ökonomischen oder vom Marketing abhängigen Gründe heute eine Abtastfrequenz von 32 kHz für ein 15-kHz-Endprodukt unzweckmäßig erscheinen lassen. Wer sollte die Verantwortung für 20 kHz tragen, wenn nach Einführung des direkten digitalen Satellitenempfangs der Hörer keinen Unterschied zwischen gekauften Tonträgern (20kHz) und dem gleichen Rundfunkprogramm-Material (15 kHz) feststellt?

44,056 kHz für NTSC und 44,1 kHz für Pal sind denkbar.

Erst nach genauer Kenntnis der Filterproblem-Lösungen kann die erforderliche Abtastfrequenz logisch rationell für die Anwendungsbereiche festgelegt werden.

Diese sind

- a) Rundfunk-/Kommunikations-Übertragungsketten (32 kHz)
- b) Fernseh-Kompatibilität mit Pal-NTSC und professionellen Audio-Recording-Studios (44 056/50,0 kHz)
- c) Vorbespielte Tonträger als Verbraucher-Endprodukt mit der Forderung nach Fidelity, Spielzeit und (Material)Preis.

Nach den vorliegenden Untersuchungsergebnissen über 15 kHz als ausreichende Bandbreite kann ohne eine sachliche Analyse der Antworten auf die bisher gestellten Fragen nicht mit einer einzigen Standard-Abtastfrequenz gerechnet werden. Mehr als zwei Frequenzen werden komplexe Probleme für Transcoder, Filter von miteinander verkoppelten Geräten (Recorder/Peripherie mit Mischpult...) erzeugen. Nicht zu unterschätzen ist auch der sachgerechte Einsatz von (unübersichtlichen) „Black-Boxes“ für Transcoder-Zwecke durch überlastetes Studio-Personal.

Digitale Echtzeit-Abtastfrequenz-Transcodierung f_x/f_y durch Interpolation eines integeren Verhältnisses (M) und einer Teilung durch das integere Verhältnis (N) erfordert heute keinen wesentlichen Mehraufwand innerhalb eines gewissen Filter-Koeffizientenbereiches.

⁷⁾ Over-Dubbing = Synchron-Aufnahmevergung zusätzlicher Tonsignale zu bereits bespielten (Takt-)Spuren eines Mehrspurbandes

⁸⁾ Aliasing-Produkte = Mischprodukte zwischen codiertem Signal und Abtastfrequenz.

⁶⁾ Codec = Coder-Decoder-Kombination.

Damit wird die Frequenzwahl erleichtert. Entsprechend den bisherigen Ausführungen sollte eine Abtastfrequenz für den professionellen Bereich, falls wirklich notwendig(!) so niedrig wie möglich gewählt werden. Der Bereich oberhalb 45 kHz erscheint zumindest langfristig gesehen als nicht erforderlich.

Auch mit Rücksicht auf einen späteren Übergang vom Videorecorder zur Longitudinal-Recordertechnik bleibt die Pal/NTSC-Kompatibilitätsforderung bestehen. Mit einem unhörbaren Tonhöhenunterschied von 0,1 % und einer für Einzelfälle externen Recorder-Synchronisierungsmöglichkeit könnte das Frequenzpaar 44,056 kHz für NTSC und 44,1 kHz (oder 44 kHz) für Pal nebeneinander bestehen.

Dynamische Quantisierungsrauschwerte (Verzerrungen) am Ende einer digitalen Übertragungskette werden grundlegend mit der Elementanzahl/Abtast-Codewort (gleichförmiger Code) bestimmt. Weitere Rauschquellen entstehen durch physikalische Grenzen bei der A/D-D/A-Wandlung. Daher müssen Minimalanforderungen für diese Wandlerstufen getrennt und auch als Codec betrachtet für die professionelle Anwendung definiert werden. Dabei hat man sich auch zu fragen:

1. Welche Geräuschmeßmethode für digitale Signalübertragung definiert Minimalwerte für Dynamik-/Frequenzabhängige Quantisierungs-Nebenprodukte?
2. Welche Minimalwerte gelten für professionelle Signalübertragung.
3. Welche Umgebungstemperatureinflüsse sind bei Wandlern zu berücksichtigen? Ist eine 14-Bit-Quantisierung mit etwa 90 dB Abhörlautstärke im Regieraum abgehört für Nachhallausklänge ausreichend?
4. Was sind im Gegensatz zur Theorie die in der Praxis erreichten Geräuschunterschiede zwischen 14- und 16-Bit-Codecs?
5. Welche (signifikant zu beweisenden) Vorteile hat ein in seinen dynamischen Eigenschaften ungenauer 16-Bit-Codec gegen-

über einer erheblich genaueren 14-Bit-Ausführung oder
6. wenn 14-Bit-Abtastwerte mit der Frequenz (x) übertragen mittels Integration aus $f_x \cdot 4$ Abtastwerten eine 16-Bit-Qualität ergeben?

Eine nicht hörbare Datenreduktion wird auch in der Studio-Technik wirtschaftliche Vorteile mit längeren Spielzeiten, verringerter Packungsdichte (Fehlerwahrscheinlichkeit) oder weniger aufwendigen Verzögerungsleitungen eingesetzt werden. Die konstante Elementmenge/Übertragungs-Codewort ist unerlässlich zum einfachen, nicht zusätzlich Fehlerquellen verursachenden Datendekodieren. Besonders von der EBU⁹⁾ sind dafür verschiedene Verfahren mit für den Konsumbereich ausreichender Qualität veröffentlicht worden.

Für die Studioteknik könnte auch das folgende Reduktionsverfahren mit gleichförmiger PCM-Übertragungsqualität vielversprechend sein: Binäre PCM-Augenblickswertdifferenzen entsprechen codierten „Slew-Rate“-Werten¹⁰⁾, die unregelmäßig verteilt und statistisch ermittelt nur selten auftreten. Kennlinien für konstante Amplitudendifferenz/Änderungsschnelle (Frequenz)-Werte entstehen, ausgehend von der maximalen Quantisierungsfähigkeit eines A/D-Wandlers bei der Abtastfrequenz (f_s) in 6 dB-Pegelstufen je reduziertem MSB¹¹⁾ bei gleichzeitig stufenweise halbiert Abtastfrequenz (f_s) (als Eckwert für die obere Frequenzgrenze, bis zu der nach dem Verfahren garantiert jeder „Slew-Rate“-Wert übertragen wird.).

Der hier skizzierte Fragenkatalog ist nur ein Rahmenprogramm zu logischen Annäherung an eine harmonisierte Digital-Studio-Technik. Die aktive Beteiligung an der Lösung einzelner Problemgruppen ist für alle am Ergebnis interessierten Firmen und Institutionen unerlässlich.

⁹⁾ EBU = European Broadcasting Union.

¹⁰⁾ Slew-Rate = Amplituden-Änderungsgeschwindigkeiten.

¹¹⁾ MSB = Most Significant Bit.

Drehknöpfe.

Formschönes Design.
Mattierte Oberfläche, Erhöhte Haltbarkeit, Spannungsbefestigung.
Vielfältige Kombinationsmöglichkeiten.
Knopfgrößen: 10mm, 11mm, 16mm, 20mm, 25mm, 38mm. Für 5 verschiedene Achsenausführungen.

2-Flügel Knebelknopf mit Strichmarken.

Steckknopf mit Mutterabdeckung.

2-Flügel Knebelknopf besonders geeignet für Stufenschalter.

Standard Drehknopf mit Zeiger und Strichmarke.

knitter-switch

8011 baldham/münchen, neue poststr. 17, postf. 100233
telefon 08106/4041, telex 05-28338